PCT/GB00/01035

NIL

The Patent Office

09/936845

INVESTOR IN PEOPLE

GB00/01035

The Patent Office
Concept House
Cardiff Road
Newport
South Wales
NP10 8QQ

PRIORITY DOCUMENT

SUBMITTED OR TRANSMITTED IN COMPLIANCE WITH RULE 17.1(a) OR (b)

(4)

I, the undersigned, being an officer duly authorised in accordance with Section 74(1) and (4) of the Deregulation & Contracting Out Act 1994, to sign and issue certificates on behalf of the Comptroller-General, hereby certify that annexed hereto is a true copy of the documents as originally filed in connection with the patent application identified therein.

In accordance with the Patents (Companies Re-registration) Rules 1982, if a company named in this certificate and any accompanying documents has re-registered under the Companies Act 1980 with the same name as that with which it was registered immediately before re-registration save for the substitution as, or inclusion as, the last part of the name of the words "public limited company" or their equivalents in Welsh, references to the name of the company in this certificate and any accompanying documents shall be treated as references to the name with which it is so re-registered.

In accordance with the rules, the words "public limited company" may be replaced by p.l.c., plc, P.L.C. or PLC.

Re-registration under the Companies Act does not constitute a new legal entity but merely subjects the company to certain additional company law rules.

Signed R. Mahoney

Dated 6th April 2000

Patents Form 1/77

Patents Act 1977
(Rule 16)

THE PATENT OFFICE
18 MAR 1999
RECEIVED BY HAND

The Patent Office

1/77

19MAR99 E433889-1 D01038
P01/7700 0.00 - 9906307.5

The Patent Office
Cardiff Road
Newport
Gwent NP9 1RH

Request for grant of a patent

(See the notes on the back of this form. You can also get an explanatory leaflet from the Patent Office to help you fill in this form)

1. Your reference — JPD/GJW/KLB/UNIBR22/43675/000

2. Patent application number *(The Patent Office will fill in this part)* — **9906307.5**

3. Full name, address and postcode of the or of each applicant *(underline all surnames)*

   UNIVERSITY OF BRISTOL
   SENATE HOUSE
   TYNDALL AVENUE
   BRISTOL
   BS8 1TH

   Patents ADP number *(if you know it)*

   If the applicant is a corporate body, give the country/state of its incorporation — 7981810001

4. Title of the invention — NOVEL POLYPEPTIDES

5. Name of your agent *(if you have one)*

   "Address for service" in the United Kingdom to which all correspondence should be sent *(including the postcode)*

   WITHERS & ROGERS
   GOLDINGS HOUSE
   2 HAYS LANE
   LONDON
   SE1 2HW

   Patents ADP number *(if you know it)* — 1776001

6. If you are declaring priority from one or more earlier patent applications, give the country and the date of filing of the or of each of these earlier applications and *(if you know it)* the or each application number

Country	Priority application number *(if you know it)*	Date of filing *(day / month / year)*

7. If this application is divided or otherwise derived from an earlier UK application, give the number and the filing date of the earlier application

Number of earlier application	Date of filing *(day / month / year)*

8. Is a statement of inventorship and of right to grant of a patent required in support of this request? *(Answer 'Yes' if:*

   a) any applicant named in part 3 is not an inventor, or
   b) there is an inventor who is not named as an applicant, or
   c) any named applicant is a corporate body.
   See note (d))

   YES

JPD/UNIBR22
O:\SP_LWP\JPD\Unibr22

Novel Polypeptides

The present invention relates to polyunsaturated fatty acid (PUFA) elongases. More specifically, the invention relates to DNA sequences from *C. elegans* encoding PUFA elongases.

The synthesis of PUFAs i.e. fatty acids of 18 carbons or more in length and containing two or more double bonds, is thought to be catalyzed in a variety of organisms by a specific fatty acid elongase enzyme. This elongase is responsible for the addition of 2 carbon units to an 18 carbon PUFA, resulting in a 20 carbon fatty acid. An example of this reaction is the elongation of γ-linolenic acid (GLA; $18:3\Delta^{6,9,12}$) to di-homo-γ-linolenic acid (DHGLA; $20:3\Delta^{8,11,14}$) in which the tri-unsaturated 18 carbon fatty acid is elongated by the addition of a two carbon unit to yield the tri-unsaturated 20 carbon fatty acid. Since there is considerable interest in the production of long chain PUFAs of more than 18 carbons in chain length, for example arachidonic acid and eicosapentanoic acid, the identification of this enzyme is of both academic and commercial interest. At present, there are no examples of identified cloned genes encoding PUFA elongases, though a number of genes encoding enzymes likely to be involved in other aspects of lipid synthesis have been identified. For example, an *Arabidopsis* gene (FAE1) has been shown to be required for the synthesis of very long chain monounsaturated fatty acids (such as erucic acid; $20:1\Delta^{11}$) (James *et al*, (1995) *Plant Cell* **7**, 309-319). However, it is clear that this enzyme does not recognize di- and tri-unsaturated 18 carbon fatty acids, for example, linoleic acid, $18:2\Delta^{9,12}$ or α-linolenic acid, $18:3\Delta^{9,12,15}$ respectively, as substrates, and is therefore not involved in the synthesis of long chain PUFAs (Millar & Kunst (1997), *Plant Journal* **12**, 121-131). This in itself is not surprising, since, of the plant kingdom, only a very few lower plant species, such as the moss *Physcomicotrella patens* (Girke *et al.*, (1998), *Plant J*, **15**: 39-48); are capable of synthesising long chain PUFAs, and therefore *Arabidopsis* would not be expected to contain any such enzymes (Napier *et al.* (1997), Biochem J, **328**: 717-720; Napier *et al.*, (1999) *Trends in Plant Sci* **4**, 2-5).

An object of the invention is to provide an isolated PUFA elongases.

Using the above-mentioned *C. elegans* genomic sequence, together with suitable search strings, the inventors identified eight related putative open reading frames (ORFs) encoding for PUFA elongases. A number of different search criteria were applied to identify a number of (ORFs) which are likely to encode polypeptides with fatty acid elongase activities.

Accordingly, a first aspect of the invention provides an isolated polypeptide comprising a functional long chain polyunsaturated fatty acid (PUFA) elongase. This polypeptide can be used to elevate PUFA levels in animals.

Preferably, the polypeptide extends the chain length of an 18 carbon PUFA to 20 carbons in length.

Preferably, the polypeptide is from an animal, more preferably, the animal is an invertebrate such as a worm. Where the animal is a worm, it is preferably *C. elegans*. Alternatively, the animal is a vertebrate, preferably a mammal such as a human, rat or mouse.

A second aspect of the invention provides an isolated DNA sequence, preferably a cDNA sequence, encoding a polypeptide according to a first aspect of the invention. This DNA sequence can be used to engineer transgenic organisms.

Preferably, the DNA sequence comprises any one of the sequences shown in SEQ ID1 to SEQ ID8, or variants of those sequences due to base substitutions, deletions, or additions.

A third aspect of the invention provides a transgenic animal engineered to express a polypeptide according to a first aspect of the invention. The transgenic animal may be engineered to express elevated levels of the polypeptide.

Preferably, the animal is a mammal such as a rat, mouse or monkey. The animal may be a lower eukaryote such as a yeast, or the animal may be a fish.

to a first aspect of the invention or a DNA sequence according to a second aspect of the invention.

Preferably, the mammal is a human.

The invention will now be described, by way of example only, with reference to SEQ ID1 to 16, and Figures 1 to 7, in which;

SEQ ID1 to 8 show the genomic DNA sequences encoding PUFA elongases A to H respectively; and

SEQ ID9 to 16 show the deduced amino acid sequences of PUFA enlongases A to H respectively; and

Figures 1 to 8 show hydrophobicity plots for each of PUFA elongases A to H respectively.

Introduction to general strategy

Initially the *C. elegans* databases were searched for any sequences which showed homology to yeast ELO genes, using the TBLASTN programme. A similar search was carried out using short (20 to 50 amino acid) stretches of ELO genes which were conserved amongst the three ELO polypeptide sequences. *C. elegans* sequences which were identified by this method were then used themselves as search probes, to identify any related *C. elegans* genes which the initial search with the yeast sequences failed to identify. This was necessary because the level of homology between the yeast ELO genes and any worm genes is always low (see BLAST scores later). To allow for a more sensitive search of worm sequences, a novel approach was adopted to circumvent the major drawback with searches using the BLAST programmes, namely that the search string (i.e. the input search motif) must be longer than 15 characters for the algorithm to work. Thus, if it was desired to search for a short motif (like a histidine box), then the BLAST programme would not be capable of doing this. A complete list of all the predicted ORFs present in the *C. elegans* genome exists as a database called Wormpep, which is freely available from the Sanger WWW site (http://www.sanger.ac.uk/Projects/C_elegans/webace_front_end.shtml). The latest version of

(ftp://ftp.sanger.ac.uk./pub/databases/wormpep) using manual search strings in MSWord 6, identified a number of *C. elegans* ORFs which contained presumptive histidine boxes. Wormpep contains predicted proteins from the *Caenorhabditis elegans* genome sequence project, which is carried out jointly by the Sanger Centre in Cambridge, UK and Genome Sequencing Center in St. Louis, USA. The current Wormpep database, Wormpep 16, contains 16,332 protein sequences (7,120,115 residues). Search strings used included [HXXHH], [HXXXHH], [QXXHH] and [YHH]. Comparison of the data from the two different searches indicated a small (<10) number of putative ORFs as candidate elongases. The histidine box motifs are shown in bold in SEQ ID 9 to 16.

Hydrophobicity plot analysis

Since the fatty acid elongase reaction is predicted to be carried out on the cytosolic face of the endomembrane system (Toke & Martin (1996), *supra*; Oh *et al* (1997), *supra)*, the putative *C. elegans* ORFs were examined for potential membrane spanning domains, via Kyte & Doolittle hydrophobicity plots *(J. Mol Biol,* (1982), **157**, 105-132). This revealed a number of ORFs with possible membrane-spanning domains, and also indicated a degree of similarity in the secondary-structure of a number of identified ORFs.

Screening for ER-retention signal sequences

The inventors postulated that since fatty acid elongases are expected to be endoplasmic reticulum (ER) membrane proteins, they might be expected to have peptide signals which are responsible for "ER-retention". In the case of ER membrane proteins, this signal often takes the form of a *C*-terminal motif [K-K-X_{2-3}-Stop], or similar variants thereof (Jackson *et al.,* (1990), *EMBO J.,* **9**, 3153-3162). Further sequence analysis of the *C. elegans* putative elongases revealed that 4 ORFs (F41H10.7, F41H10.8, F56H11.4, Y53F4B.c) had *C*-terminal motifs that exactly matched this search pattern, and that a further 2 ORFs (F11E6.5, C40H1.4) had related sequences. These sequence motifs are underlined in SEQ ID 9 to 13, 15 and 16.

Chromosome mapping

Since the inventors had previously observed that *C.elegans* genes involved in the synthesis of PUFA may exist in tandem (for example the Δ5 and Δ6 desaturases required for AA and

G	F56H11.4#	Z68749	IV, 2.5
H	Y53F4B.c	Z92860	II

* or # indicates genes in tandem

Comparison of *C. elegans* putative elongase ORFs with yeast genes:

Each of the three yeast ELO polypeptides were compared against all of the worm putative elongase translated ORF sequences, and then ranked in order of similarity (as measured by the BLAST score) (Altschul *et al* (1990), *supra*)

The results are shown below, with the ORF sequences ranked from most similar to least similar, and the BLAST scores are shown in brackets:

Yeast ELO1 (14 to 16 carbon fatty acid elongase)

G (262) > E (241) > D (225) > C (219) > A (216) > F (215) > H (197) > B (172)

Yeast ELO2 (24 carbon sphingolipid elongase)

E (231) > C (226) > G (189) > A (181) > F (166) > D (150) > H (141) > B (140)

Yeast ELO3 (24 to 26 sphingolipid elongase)

D (171) > G (163) > F (154) > A (152) > E (150) > C (131) > B (132) > H (128)

It is clear from the numeric values of the BLAST scores that the sequences are related, but the levels of homology are low. For comparison, the BLAST score for homology between two related worm proteins, the Δ5 and the Δ6 desaturase is in excess of 500.

Expression f *C. elegans* elongase in plants

In order to express *C. elegans* elongase in plants, the following protocol can be used to create the transgenic plants. *C. elegans* ORF sequence can be subcloned into a plant expression vector pJD330, which comprises a viral 35S promoter, and a Nos terminator. The resulting cassette or promoter/coding sequence/terminator can then be subcloned into the plant binary transformation vector pBin 19, and the resulting plasmid introduced into *Agrobacterium tumefaciens.* This *Agrobacterium* strain can then be used to transform Arabidopsis by the vacuum-infiltration of inflorescences, and the seeds harvested and plated onto selective media containing kanamycin. Since pBin 19 confers resistance to this antibotic, only transformed plant material will grow. Resistant lines can therefore be identified and self-fertilized to produce homozygous material. Leaf material can then be analyzed for expression of *C. elegans* elongase.

F11E6.5

atggcagcagcacaaacaagtccagcagccacgctcgtcgatgttttgacaaaaccatgg
agtctggatcagactgattcttacatgtctacatttgtaccattatcctataaaatcatg
attggttatctcgtcaccatctacttcgggcaaaaattaatggctcacagaaaaccattc
gatctccaaaatacacttgctctctggaacttcgggttttcactgttctcgggaatcgcc
gcctataagcttattccagaactattcggagttttcatgaaggacgggtttgtcgcttcc
tactgtcaaaacgagaactactacaccgatgcatcaactggattctggggctgggccttt
gtgatgtcgaaagctccagaactaggggatactatgttcttggtccttcgtaaaaaacca
gttatcttcatgcactggtatcatcatgccctcacatttgtctacgcagtagtcacatac
tctgagcatcaggcatgggctcgttggtctttggctctcaaccttgccgtccacactgtt
atgtatttctacttcgccgttcgcgccttgaacatccaaactccacgcccagtggcaaag
ttcatcactactattcaaattgtccaatttgtcatctcatgctacatttttgggcatttg
gtattcattaagtctgctgattctgttcctggttgcgctgttagctggaatgtgctatcg
atcggaggactcatgtacatcagttatttgttccttttgccaagttcttctacaaggcc
tacattcaaaaacgctcaccaaccaaaaccagcaagcaggagtag

SEQ ID4

F41H10.7

atgtcatcggacgatcgtggcactagaaccttcaagatgatggatcaaattcttggaaca
aacttcacttatgaaggtgccaaagaagttgctcgaggccttgaaggtttctcagcaaag
cttgccgtcggatatattgccactatttttggactgaaatattatatgaaagaccgaaaa
gccttcgatctcagtactccattaaacatttggaatggtattctttcgacattcagctta
ttgggattcttattcacttttcctactttgttatcagttatcagaaaggatggatttagt
cacacctattcccatgtctctgagctttacactgacagtacctctggatattggatcttc
ctttgggttatctcaaagattccggaacttttggatacagtattcattgttcttcgcaag
agaccacttattttcatgcactggtaccatcacgcattgaccggttactatgctcttgtc
tgctaccatgaggatgctgtccatatggtttgggttgtatggatgaattatattattcat
gcattcatgtatggatactatcttctgaaatctctgaaagttccaattccaccatcagtt
gctcaagcaatcaccacatctcaaatggttcaattcgcagttgccattttcgcacaagtt
catgtttcctataaacactatgttgagggagttgaaggattagcctactcgttcagagga
acagctatcggatttttcatgcttactacctacttctatctatggattcaattctacaaa
gagcactatcttaagaatggaggcaaaaagtacaatttggcaaaggatcaggcaaaaact
caaacaaagaaggctaactaa

SEQ ID5

F41H10.8 (ce477)

atgccacagg gagaagtctc attctttgag gtgctgacaa ctgctccatt
cagtcatgag ctctcaaaaa agcatattgc acagactcag tatgctgctt
tctggatctc aatggcatat gttgtcgtta tttttgggct caaggctgtc
atgacaaacc gaaaaccatt tgatctcacg ggaccactga atctctggaa
tgcgggtctt gctatttttct caactctcgg atcacttgcc actacatttg
gacttctcca cgagttcttc agccgtggat ttttcgaatc ttacattcac
atcggagact tttataatgg actttctgga atgttcacat ggcttttcgt
tctctcaaaa gttgctgaat tcggagatac acttttttatt attcttcgta

gccaactgt gatttcgagc catcagtatt caagctcgca
gttttcatgg acacaacata cttggctctt ttcgtcaact tcttcctcca
atcatatgtt
ctccgcggag gaaaagacaa gtacaaggca gtgccaaaga agaagaacaa ctaa

SEQ ID8

Y53F4B.c

atgtcggccg aagtgtccga acgattcaaa gtttggacag gaaacaatga
gaccatcatc tattccccat tcgagtacga ttccacgttg ctcatcgagt
catgtcggtg tacttatcag ctgcttatat tattgcgaca aatttattac
agagatatat ggagtcacgg aaacctaaaa cttttactag catggaacgg
ttttttggca gtgttcagta ttatgggtac atggagattt ggaatcgaat
tctacgatgc tgttttcaga agaggcttca tcgattcgat ctgcctggct
gtaaatccac gttcaccgtc cgcattctgg gcatgcatgt tcgctctatc
gaaaatcgcc gagtttgggg acacgatgtt cttggtgctg aggaaacggc
cggttatatt ccttcactgg tatcatcacg ctgttgttct gatcctttct
tggcatgctg caatcgaact cacagctcca ggacgctggt ttatttttat
gaactatttg gtgcattcaa taatgtatac atactacgca ataacatcaa
tcggctatcg tcttcccaaa atcgtttcaa tgactgttac attccttcaa
actcttcaaa tgctcattgg tgtcagcatt tcttgcattg tgctttattt
gaagcttaat ggagagatgt gccaacaatc ctacgacaat ctggcgttga
gcttcggaat ctacgcctca ttcctggtgc tattctccag tttcttcaac
aatgcatatt tggtaaaaaa ggacaagaaa cccgatgtga agaaggatta
a

SEQ ID9

A

1 MELAEFWNDL NTFTIYGPNH TDMTTKYKYS YHFPGEQVAD PQYWTILFQK
51 YWYHSITISV LYFILIKVIQ KFMENRKPFT LKYPLILWNG ALAAFSIIAT
101 LRFSIDPLRS LYAEGFYKTL CYSCNPTDVA AFWSFAFALS KIVELGDTMF
151 IILRKRPLIF L**HYYHH**AAVL IYTVHSGAEH TAAGRFYILM NYFAHSLMYT
201 YYTVSAMGYR LPKWVSMTVT TVQTTQMLAG VGITWMVYKV KTEYKLPCQQ
251 SVANLYLAFV IYVTFAILFI QFFVKAYIIK SSKKSKSVKN E*

SEQ ID10

B

1 MAKYDYNPKY GLENYSIFLP FETSFDAFRS TTWMQNHWYQ SITASVVYVA
51 VIFTGKKVVL IYKKSRVITF ESSLQNAIKN RNRKSLNSSQ MFQIMEKYKP
101 FQLDTPLFVW NSFLAIFSIL GFLRMTPEFV WSWSAEGNSF KYSICHSSYA
151 QGVTGFWTEQ FAMSKLFELI DTIFIVLRKR PLIFL**HWYHH** VTVMIYTWHA
201 YKDHTASGRW FIWMNYGVHA LMYSYYALRS LKFRLPKQMA MVVTTLQLAQ
251 MVMGVIIGVT VYRIKSSGEY CQQTWDNLGL CFGVYFTYFL LFANFFYHAY

201 VGVIVNLFVH AFMYPYYFTR SMNIKVPAKI SMAVTVLQLT QFMCFIYGCT
251 LMYYSLATNQ ARYPSNTPAT LQCLSYTLHL L*

SEQ ID15

G

MAQHPLVQRL LDVKFDTKRF VAIATHGPKN FPDAEGRKFF ADHFDVTIQA
SILYMVVVFG TKWFMRNRQP FQLTIPLNIW NFILAAFSIA GAVKMTPEFF
GTIANKGIVA SYCKVFDFTK GENGYWVWLF MASKLFELVD TIFLVLRKRP
LMFL**HWYHH**I LTMIYAWYSH PLTPGFNRYG IYLNFVVHAF MYSYYFLRSM
KIRVPGFIAQ AITSLQIVQF IISCAVLAHL GYLMHFTNAN CDFEPSVFKL
AVFMDTTYLA LFVNFFLQSY VLRGGKDKYK AVPKKKNN*

SEQ ID16

H

MSAEVSERFKVWTGNNETIIYSPFEYDSTLLIESCRCTYQLLILLRQI
YYRDIWSHGNLKACDXLLLAWNGFLAVFSIMGTWRFGIEFYDAVFRXG
FIXSICLAVNPRSPSAFWACMFALSKIAEFGDTMFLVLRKRPVIFL**HWYHH**
AVVLILSWHAAIELTAPGRWFIFMNYLVHSIMYTYYAITSIGYRXPKIVSMT
VTFLQTLQMLIGVSISCIVLYLKLNGEMCQQSYDNLALSFGIYASFLVLSSFF
NNAYLVKKDKKPDVKKD*

15. A transgenic animal according to claim 14 wherein the mammal is a rat, mouse or monkey.

16. A transgenic animal according to claim 13 wherein the animal is a lower eukaryote.

17. A transgenic animal according to claim 16 wherein the lower eukaryote is a yeast.

18. A transgenic animal according to claim 13 wherein the animal is a fish.

19. A transgenic plant engineered to express a polypeptide according to any of claims 1 to 10.

20. A PUFA produced by a reaction catalysed by a polypeptide according to any of claims 1 to 10.

21. A PUFA according to claim 20 wherein the PUFA is di-homo-gamma-linoleic acid ($20:3\Delta^{8,11,14}$), arachidonic acid ($20:4\Delta^{5,8,11,14}$), eicosapentanoic acid ($20:5\Delta^{5,8,11,14,17}$), docosatrienoic acid ($22:3\Delta^{3,16,19}$), docosatetraenoic acid ($22:4\Delta^{7,10,13,16}$), docosapentaenoic acid ($22:5\Delta^{7,10,13,16,19}$) or docosahexaenoic acid ($22:6\Delta^{4,7,10,13,16,19}$).

22. A PUFA according to claim 20 wherein the PUFA is a 24 carbon fatty acid with at least 4 double bonds.

23. A foodstuff comprising a PUFA according to any of claims 20 to 22.

24. A dietary supplement comprising a PUFA according to any of claims 20 to 22.

25. A pharmaceutical composition comprising a polypeptide according to any of claims 1 to 10.

26. A pharmaceutical composition comprising a PUFA according to any of claims 20 to 22.

27. A pharmaceutical composition according to claim 25 or claim 26 wherein the composition comprises a pharmaceutically-acceptable diluent, carrier, excipient or extender.

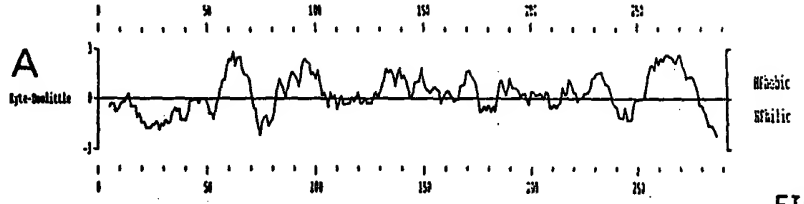

FIG 1

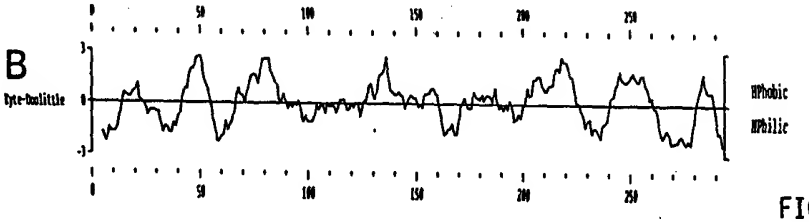

FIG 2

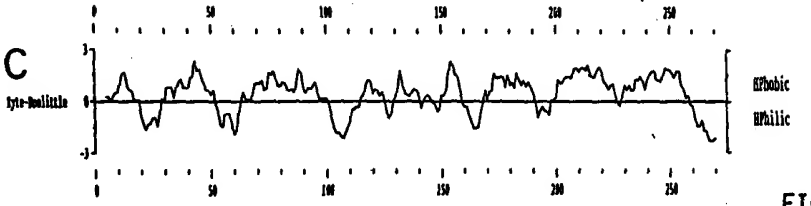

FIG 3

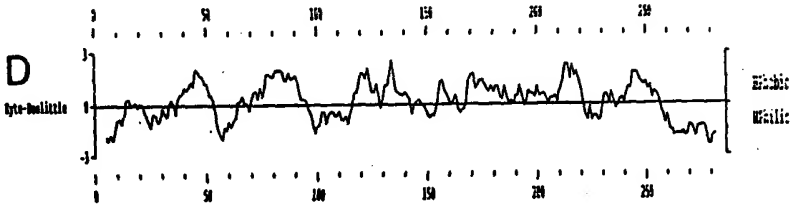

FIG 4

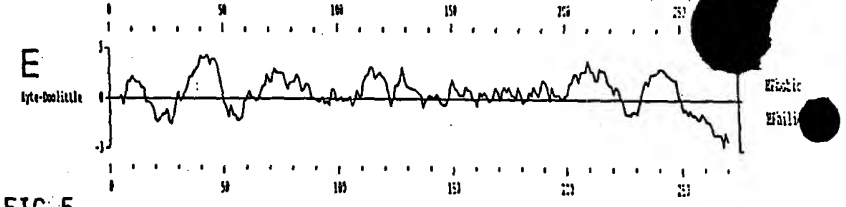

FIG 5

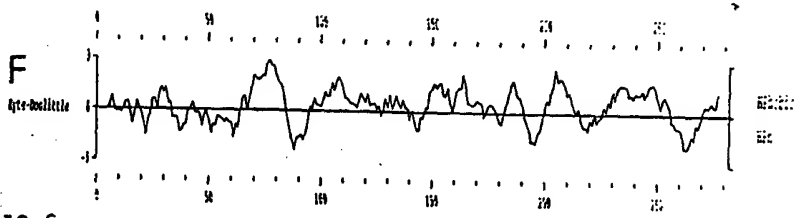

FIG 6

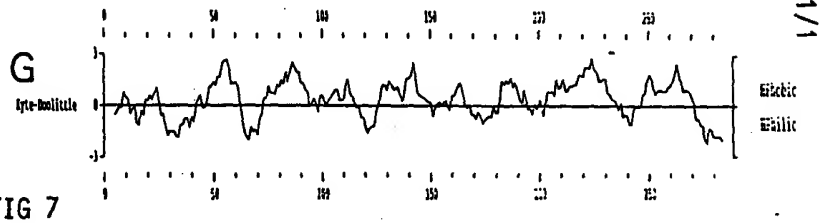

FIG 7

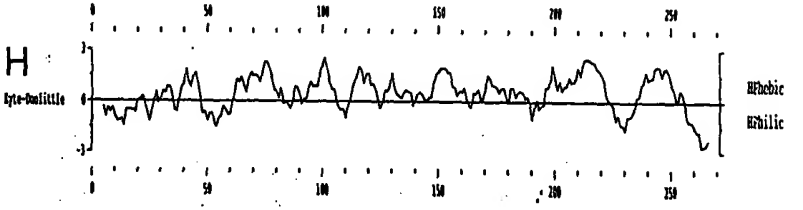

FIG 8

1/1